224. *The Crystalline Structure of the Sugars. Part I. Simple Sugars and Glycosides: Preliminary Data.*

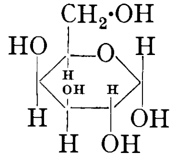
By E. G. Cox, T. H. Goodwin, and (Miss) A. I. Wagstaff.

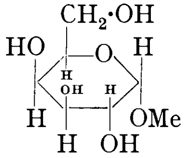
The sugars and their derivatives, on account of their importance individually and as principal components of many naturally occurring substances, merit close study by *X*-ray methods. The geometrical structure of the molecules, determinable by such methods, has considerable bearing on many of their chemical and physical properties, particularly of polymerides such as starch and cellulose, and in regard to many of the derivatives (*e.g.*, acetone compounds and anhydrides) it is of great interest from the point of view of the strain theory. On the other hand, details of the arrangement of the molecules in the crystal may be expected to contribute materially to our knowledge of the nature and extent of co-ordination or dipole association between hydroxyl and other groups in organic molecules.

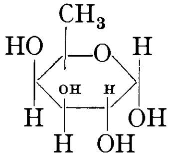
Chiefly owing to the researches of Haworth, Hirst, and their collaborators, the chemical structures of most of the sugars and their derivatives have been established beyond the possibility of doubt, but even when various assumptions are made regarding valency angles and interatomic distances, it is found that any one chemical structure comprises several geometrical possibilities. This fact, combined with the lack of any assistance from optical data, and the asymmetry of the molecules concerned, intensifies the difficulties of determining even the approximate molecular arrangement in any particular crystalline compound, while, as is well known, the exact determination by Fourier methods cannot be usefully attempted until the approximate structure has been found. Hence, although partially successful attempts have been made to suggest the molecular arrangements in various crystalline sugars and their derivatives, no great advance has been made, and the details of the structures which have been proposed, *e.g.*, for cellulose must necessarily be regarded as speculative until the geometry of the simpler carbohydrates is better known.

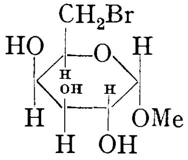
In this series of papers, we propose, first, to make a comparative survey of the cell dimensions and other data which are now available for over 60 compounds: in this way certain important general conclusions are obtained which can be applied in special cases to determine approximate structures. More detailed examinations of these compounds will be reported later, but the present paper records preliminary data for various substances which had not previously been examined, *viz.*, sorbose (the constitution of which is not yet definitely known), α-galactose (I), α-methylgalactoside (II), α-fucose (III), α-methylgalactoside-6-bromohydrin (IV), β-methylarabinoside (V), and α-methylfucoside (VI). Details for glucoheptose, lyxose, allose, altrose, ribose, and glucoheptulose (the last four made available by the generosity of Prof. W. C. Austin) will be published later. The formula for α-*d*-galactose (I) is written with the correct orientation, whereas the remaining formulæ, in order to emphasise their relationship to (I), are all oriented in the same way, irrespective of their relative configuration; thus (III) strictly represents α-*d*-fucose, although the substance actually examined was α-*l*-fucose. This is a matter of indifference from the crystallographic point of view, since it is impossible to distinguish between *d*-

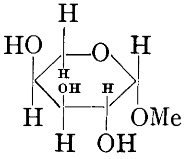
and *l*-forms by *X*-ray methods. It should be noted, further, that since the nomenclature of α- and β-forms depends upon the configuration of the *last asymmetric* carbon atom, pentoses do not necessarily correspond with methylpentoses and hexoses in this respect; thus, β-arabinose is assigned the same configuration as α-fucose and α-galactose.

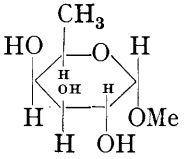

(I.)


(II.)


(III.)


(IV.)


(V.)


(VI.)

X-Ray data have previously been published for a number of simple sugars, *viz.*, β-arabinose, α-xylose, and α-rhamnose hydrate (Andress and Reinhardt, *Z. Krist.*, 1931, **78**, 477; Cox, J., 1931, 2313; *Z. Krist.*, 1932, **84**, 45), β-fructose and α-glucose (Hengstenberg and Mark, *ibid.*, 1929, **72**, 301), α- and β-glucose and α-glucose hydrate (Sponsler and Dore, *J. Amer. Chem. Soc.*, 1931, **53**, 1639; powder photographs only) and β-mannose (Marwick, *Proc. Roy. Soc.*, 1931, *A*, **131**, 621; cf. McCrea, *Proc. Roy. Soc. Edin.*, 1931, **51**, 190, who erroneously described this substance as α-mannose). Various methylglycosides have also been studied, *viz.*, α- and β-arabinosides, α- and β-glucosides (Cox and Goodwin, *Z. Krist.*, 1933, **85**, 462), α-mannoside (pyranose and furanose forms) (*idem*, J., 1932, 1844), β- and α-xylosides (Cox, J., 1932, 138, 2535) and β-rhamnoside (Braekken, Koren, and Sörensen, *Z. Krist.*, 1934, **88**, 205). From a chemical point of view, the hydroxyl attached to the first carbon atom in a sugar (the reducing group) can be sharply differentiated from the remaining hydroxyl groups. It is therefore of interest to study the crystalline methylglycosides, in which this group only has been replaced, to determine whether any marked change in structure of a general nature can be detected. A full discussion will be given at a later stage, but from a review of the material now available it already appears highly probable that the first carbon atom and its attached group possess special significance from a geometrical point of view also.

Noteworthy among the present results is the excellent example of morphotropism presented by β-methylarabinoside (V), α-methylfucoside (VI), and α-methylgalactoside-6-bromohydrin (IV). Morphotropic relationships are rare among carbohydrates, since their crystalline structures are less dependent upon the mere size and shape of the molecules than those of most organic compounds; their hardness, high densities, and high melting points suggest that the crystal structure of a sugar is largely determined by the hydroxyl groups, which tend to associate to the greatest possible extent, and link the molecules in all directions in the lattice, so that the replacement of one group in the molecule by another of different size usually necessitates a complete readjustment of the relative positions of the hydroxyl groups and therefore of the molecules themselves. In the present case, however, it appears that the hydroxylic linking is two-dimensional only, the molecules being linked together in sheets, the thickness of which can be changed appreciably without otherwise altering the structure. The cell dimensions (in Å.) of the three substances are as follows:

	Arabinoside (V).	Fucoside (VI).	Bromohydrin (IV).
d_{100}	8·10	9·96	10·58
b	7·74	7·87	7·81
c	5·89	5·72	2 × 5·62

These three substances differ only in the groups attached to the fifth carbon atom; the above figures show that the introduction of methyl (VI) or bromomethyl (IV) in place of hydrogen in (V) is accompanied only by an expansion of the *a* spacing, the other lattice dimensions remaining nearly constant; moreover, reference to a model shows that the dimensions of the molecules (an interval of 3·5 Å. being allowed between the carbon atoms of neighbouring molecules) in the direction joining the first and fifth carbon atoms correspond closely with the above values of d_{100}. Since in the arabinoside the first carbon atom carries the methoxyl group, and the fifth only hydrogen atoms, it is to be expected that the forces across the (100) plane will be relatively small, and in agreement with this it is found that the crystals exhibit good cleavage parallel to {100}. The replacement of one of the C_5 hydrogens by methyl or bromomethyl will reduce still further the cohesion across (100), and actually in (VI) and (IV) the cleavage parallel to this plane is perfect. Further implications of these results will be considered later in conjunction with other data, but it may be pointed out here that this close relation between the α-fucoside, the β-arabinoside, and the α-galactoside-6-bromohydrin constitutes very strong confirmation for the configurations which have been assigned to the groups on the first carbon atoms of these glycosides; there appears to be no correspondence whatever between, *e.g.*, the crystalline structures of α-methylfucoside and α-methylarabinoside (Cox and Goodwin, *loc. cit.*). Since criticism has been directed against deductions from rotational data, upon which these configurations are chiefly based, it is of especial interest to obtain independent crystallographic evidence of this nature. It may be noted, finally, that, whereas when one hydrogen atom on the sixth carbon atom of the fucoside is replaced by bromine to give (formally) the α-methylgalactoside-6-bromohydrin (IV) the same type of structure persists, yet the introduction of the hydroxyl group (which is actually smaller than the bromine atom) to form α-methylgalactoside produces an entirely different lattice containing water of crystallisation. This is in agreement with the view expressed above regarding the special significance of the hydroxyl group in carbohydrate crystal structures.

EXPERIMENTAL.

The cell dimensions and space-groups recorded in this and subsequent papers have been determined (unless otherwise stated) by means of single-crystal rotation and oscillation photographs, with copper K_a radiation. The probable error in the cell dimensions is about ⅓%. Densities have been determined (to 1%) by suspension in suitable liquids, and refractive indices have been measured to 0·002 by immersion methods, using sodium light.

α-d-*Galactose*.—Crystals were prepared by the method of Riiber, Minsaas, and Lyche (J., 1929, 2173), *viz.*, by pouring a concentrated aqueous solution into boiling alcohol and allowing it to cool. The initial specific rotation (+ 140°) of a sample prepared in this way showed the amount of β-galactose present to be very small. Rapid precipitation gives characteristic thin plates similar to those previously described by Ost (*Z. anal. Chem.*, 1890, **29**, 651); these plates show diagonal extinction with refractive indices 1·569 and 1·530, the latter parallel to the shorter diagonal. By slow cooling, larger prismatic crystals are obtained showing the orthorhombic combinations *r*{101}, *m*{210}, and sometimes *s*{201}; (*r* : *r'* = 62° 26', *m* : *m'* = 78° 21'). The cell dimensions are a = 12·68, b = 7·78, c = 7·71 Å., whence the axial ratios are $a : b : c$ = 1·630 : 1 : 0·991. The halvings were found to be (*h*00), (0*k*0) and (00*l*) halved for *h*, *k*, and *l* odd respectively, from which it follows that the space-group is $P2_12_12_1$ (Q^4). With four asymmetric molecules in the unit cell *d* (calc.) = 1·58 g./c.c. (obs., 1·58). The refractive indices are α = 1·530, β = 1·56 (approx.), and γ = 1·569. The optic axial angle is large, *a* being the acute bisectrix and (001) the plane of the optic axes. The optic sign is negative.

When the above measurements had been completed, *X*-ray data for α-galactose were published by Braekken, Koren, and Sörensen (*loc. cit.*), who found cell dimensions in good agreement with our results.

d-*Sorbose*.—Goniometric data for sorbose were published by Berthelot (*Compt. rend.*, 1857, **45**, 268) and by Pelouze (*Annalen*, 1852, **83**, 47), who described it as orthorhombic with $a : b : c$ = 0·3357 : 1 : 0·3523. The density is recorded as 1·654 (Berthelot) or 1·612 (de Bruyn and van Ekenstein). The specimen used for the present study was prepared (by Mr. J. F. Preston) by the action of *B. xylinum* on sorbitol, and recrystallised from aqueous alcohol. The crystals were orthorhombic combinations of *m*{110} and *r*{101}, usually equally developed, but sometimes elongated along the *c*-axis, and occasionally showing the form {130} (*m* : *m'* = 39° 45', *r* : *r'* =

38° 20′). The cell dimensions are $a = 18{\cdot}01$, $b = 6{\cdot}51$, $c = 6{\cdot}26$ Å., the axial ratios being $a : b : c = 2{\cdot}766 : 1 : 0{\cdot}962$. The rearrangement of these ratios in the form $c : a : b = 0{\cdot}347 : 1 : 0{\cdot}361$ shows that the crystals used by the earlier investigators were essentially the same as those now examined. The space-group is $P2_12_12_1$ (Q^4), and with four asymmetric molecules in the unit cell, d (calc.) = 1·63 g./c.c. (obs., 1·63). The crystals are optically negative, the refractive indices being $\alpha = 1{\cdot}553$ (parallel to a), $\beta = 1{\cdot}566$ (parallel to c), and $\gamma = 1{\cdot}572$.

α-1-*Fucose*.—We are indebted to Mr. J. Minsaas of Trondheim for the loan of a specimen of this sugar ($[\alpha]_D - 158°$). After numerous attempts at recrystallisation, crystals of sufficient size for *X*-ray examination were obtained from absolute alcohol, but these were so poorly developed that the cell dimensions recorded below are accurate only to about 1%. The crystals are orthorhombic prisms *m*{011} terminated by *q*{101}. The cell dimensions are $a = 14{\cdot}3$, $b = 7{\cdot}6$, and $c = 6{\cdot}6$ Å., and the space-group is $P2_12_12$ (Q^3). Although {001} is not halved, the first and third orders are very weak. With four asymmetric molecules in the unit cell, d (calc.) = 1·52 g./c.c. (obs., 1·49). The minimum and the maximum refractive index are 1·537 and 1·557.

β-*Methyl*-1-*arabinoside*.—This substance (m. p. 170°, $[\alpha]_D + 245°$) is dimorphous, data for an orthorhombic form having previously been published (Cox and Goodwin, *Z. Krist.*, 1933, **85**, 462). The monoclinic variety now described exhibits the forms *a*{100}, *m*{110}, and *c*{001} (*a* : *m* = 46° 21′); the crystals are elongated in the *c*-direction and show good cleavage parallel to *a*{100}, which is usually the most prominent form. The cell dimensions are $a = 8{\cdot}99$, $b = 7{\cdot}74$, $c = 5{\cdot}89$ Å.; $\beta = 115°\ 35'$, and since {010} is halved, the space-group is $P2_1$ (C_2^2). With two asymmetric molecules in the unit cell, d (calc.) = 1·47 g./c.c. (obs., 1·46). The refractive indices are $\alpha = 1{\cdot}524$, $\beta = 1{\cdot}529$, $\gamma = 1{\cdot}546$. The optic axial angle is $2V = 71°$, c being the optic normal and b the acute bisectrix; a centred obtuse bisectrix figure is visible through *a*{100}. Since this work was completed Braekken, Koren, and Sörensen (*loc. cit.*) have published cell dimensions agreeing closely with those above.

α-*Methyl*-1-*fucoside*.—We are indebted to Mr. J. Minsaas for a loan of a specimen of this substance (m. p. 158°, $[\alpha]_D - 197°$). Recrystallisation from ethyl acetate yielded moderately good monoclinic sphenoidal crystals, tabular on *a*{100}, and exhibiting also the forms *m*{110}, *R*$\{10\bar{1}\}$, *o*{111}, and *O*$\{1\bar{1}\bar{1}\}$. The forms *r*{101}, *q*{011}, *ω*$\{1\bar{1}1\}$, and Ω$\{11\bar{1}\}$ were occasionally observed. Cleavage parallel to *a*{100} is perfect. In the following table the principal interfacial angles determined goniometrically are compared with those calculated from the *X*-ray data.

	Calc.	Obs.		Calc.	Obs.
a(100) : *m*(110)	51° 16′	51° 05′	*a*(100) : *O*$(1\bar{1}\bar{1})$	106° 12′	105° 40′
a(100) : *o*(111)	56 53	56 54	*a*(100) : *q*(011)	79 47	79 38
a(100) : *r*(101)	51 08	50 45	*o*(111) : *m*(110)	43 29	43 30
a(100) : *R*$(10\bar{1})$	70 15	69 37	*m*(110) : *ω*$(1\bar{1}1)$	92 25	91 57

The classification angles (T. V. Barker's system) are : *a*(100) : *r*(101) = 51° 08′; *r*(101) : *c*(001) = 26° 16′; *a*(100) : *m*(110) = 51° 16′; *b*(010) : *q*(011) = 54° 41′. The cell dimensions are $a = 10{\cdot}06$, $b = 7{\cdot}87$, and $c = 5{\cdot}72$ Å., whence $a : b : c = 1{\cdot}277 : 1 : 0{\cdot}726$; $\beta = 102°\ 36'$. The space-group is $P2_1$, and with two asymmetric molecules in the unit cell, d (calc.) = 1·33 g./c.c. (obs., 1·31).

The two refractive indices in the plane (100) are 1·488 parallel to c and $\gamma = 1{\cdot}513$ (parallel to b). The index perpendicular to (100) could not be determined accurately, but apparently neither α nor β is greatly different from 1·49, b being the acute bisectrix. The optic orientation is thus similar to that of β-methylarabinoside.

α-*Methyl*-d-*galactoside Monohydrate*.—This substance (m. p. 110°, $[\alpha]_D + 179°$) was studied by Reuter (*Z. Krist.*, 1902, **35**, 389) who described it as orthorhombic bisphenoidal with axial ratios $a : b : c = 0{\cdot}6225 : 1 : 1{\cdot}7418$, the forms observed being *r*{101} and *q*{011} equally developed, with *s*{021} and *c*{001} less prominent. By a misprint it is recorded in Groth ("Chemische Krystallographie," **3**, 447) as the racemic form. The crystals used for the present investigation corresponded with Reuter's description, but in accordance with the usual convention we have re-allotted the axes in the order $a > b > c$. The cell dimensions are $a = 21{\cdot}33$, $b = 7{\cdot}45$, and $c = 6{\cdot}12$ Å., whence $a : b : c = 2{\cdot}862 : 1 : 0{\cdot}824$, Reuter's recalculated figures being $c : a : b = 2{\cdot}798 : 1 : 2 \times 0{\cdot}824$. The space-group is $P2_12_12_1$, and with four asymmetric molecules in the cell d (calc.) = 1·44 g./c.c. The refractive indices are $\alpha = 1{\cdot}522$ (parallel to c), $\beta = 1{\cdot}523$ (parallel to b), and $\gamma = 1{\cdot}529$ (Reuter, *loc. cit.*).

α-Methylgalactoside-6-*bromohydrin.*—A specimen of this substance ($[\alpha]_D = +157°$) was obtained through the generosity of Dr. F. Valentin, of Prague, who also communicated to us the unpublished crystallographic results of Mons. R. Novacek. This worker found the crystals to be orthorhombic bisphenoidal combinations of *b*{010}, *c*{001}, *o*{111}, and occasionally *a*{100} and *O*{11$\bar{1}$}; they are elongated along the *a* axis and exhibit perfect cleavage parallel to *b*{010}. Although the crystals are large and clear, their faces are imperfect, and the axial ratios were therefore determined only approximately; $a : b : c = 0{\cdot}771 : 1 : 1{\cdot}051$. The refractive indices were found to be $\alpha = 1{\cdot}581$, $\beta = 1{\cdot}584$, and $\gamma = 1{\cdot}586$, the plane of the optic axes being *c*(001) and the *a* axis the acute bisectrix.

In accordance with convention, Novacek's *a* and *c* axes should be interchanged; in order to bring out the parallelism between this substance and β-methylarabinoside, however, we have interchanged his *a* and *b* axes, so that the axial ratios become $a : b : c = 1{\cdot}297 : 1 : 1{\cdot}363$, and the cleavage plane becomes *a*{100}. The cell dimensions are $a = 10{\cdot}58$, $b = 7{\cdot}81$, and $c = 11{\cdot}23$ Å., whence $a : b : c = 1{\cdot}354 : 1 : 1{\cdot}438$, with which the goniometric measurements, in view of their approximate nature, are in reasonably good agreement. The space-group is $P22_12_1$ (Q^3), and with four asymmetric molecules in the unit cell, *d* (calc.) = 1·84 g./c.c. (obs., 1·86).

The authors are indebted to the Department of Scientific and Industrial Research for a Senior Research Award to one of them (T. H. G.), and to Prof. W. N. Haworth, F.R.S., for his interest in the work.

UNIVERSITY OF BIRMINGHAM.

[*Received, June 4th*, 1935.]